

## APPLYING AI TECHNOLOGY IN PROTECTION OF SWIFTLETS FROM SOME PREDATORS

Luu Trong Hieu<sup>1</sup>, Tran Anh Dung<sup>2</sup>, Nguyen Chi Ngon<sup>1\*</sup>

<sup>1</sup>Can Tho University, <sup>2</sup>An Giang Vocational College

ARTICLE INFO	ABSTRACT
<p><b>Received:</b> 09/7/2021</p> <p><b>Revised:</b> 12/8/2021</p> <p><b>Published:</b> 18/8/2021</p>	<p>In recent years, swiftlet farming is developing rapidly with the ability to get high profits and provide many jobs for farmers. However, some popular predators such as snakes, hawks and owls, are a big problem that cannot be ignored. This paper proposes a model of repelling and alarming when the predatory animals enter the swiftlets' house. A deep learning model-based software is developed to identify some popular predators from images captured by camera. An IoT-based hardware is applied to send alarm messages to mobile phone when the software detects snake attacked or turn on a light to repel hawks and owls. First stage testing results show that the deep learning model can classify some popular predators with over 90% accuracy while the hardware system can text to right phone numbers when detecting snakes, and turn on the light when recognizing hawks or owls.</p>
<p><b>KEYWORDS</b></p> <p>GoogLeNet</p> <p>Deep learning</p> <p>Classification</p> <p>Identification</p> <p>IoT</p>	

## ỨNG DỤNG CÔNG NGHỆ AI TRONG BẢO VỆ NHÀ YẾN TỪ ĐỘNG VẬT SĂN MỒI

Luu Trọng Hiếu<sup>1</sup>, Trần Anh Dũng<sup>2</sup>, Nguyễn Chí Ngôn<sup>1\*</sup>

<sup>1</sup>Trường Đại học Cần Thơ,

<sup>2</sup>Trường Cao đẳng nghề An Giang

THÔNG TIN BÀI BÁO	TÓM TẮT
<p><b>Ngày nhận bài:</b> 09/7/2021</p> <p><b>Ngày hoàn thiện:</b> 12/8/2021</p> <p><b>Ngày đăng:</b> 18/8/2021</p>	<p>Những năm gần đây, nghề nuôi chim yến phát triển nhanh chóng với khả năng tạo ra lợi nhuận cao và giải quyết nhiều việc làm cho người dân. Tuy nhiên, một số động vật ăn thịt phổ biến như rắn, điều hâu và cú, là một vấn đề lớn không thể bỏ qua. Bài báo này đề xuất một mô hình xua đuổi và báo động khi động vật săn mồi xâm nhập vào nhà yến. Một phần mềm dựa trên mô hình học sâu được phát triển để xác định một số loài săn mồi phổ biến từ hình ảnh được camera chụp lại. Một phần cứng dựa trên nền tảng IoT được áp dụng để gửi tin nhắn cảnh báo đến điện thoại di động khi phần mềm phát hiện rắn tấn công hoặc bật đèn để xua đuổi điều hâu và cú. Kết quả thử nghiệm giai đoạn đầu cho thấy, mô hình học sâu có thể phân loại một số loài săn mồi phổ biến với độ chính xác trên 90%, trong khi hệ thống phần cứng có thể nhắn tin đến đúng số điện thoại khi phát hiện rắn và bật đèn khi nhận ra điều hâu hoặc cú.</p>
<p><b>TỪ KHÓA</b></p> <p>Mạng GoogLeNet</p> <p>Phương pháp học sâu</p> <p>Phân loại đối tượng</p> <p>Nhận dạng đối tượng</p> <p>Internet vạn vật</p>	

DOI: <https://doi.org/10.34238/tnu-jst.4736>

\* Corresponding author. Email: ncngon@ctu.edu.vn

## 1. Giới thiệu

Những năm gần đây, nghề nuôi chim yến ở Việt Nam phát triển mạnh do mang lại nhiều lợi ích về kinh tế và giải quyết việc làm về mặt xã hội cho nhiều người dân. Nhà nuôi yến là nơi để dẫn dụ chim yến về sống và làm tổ. Cũng như các hang yến ngoài tự nhiên, nhà yến bị khá nhiều loài gây hại xâm nhập. Các loài gây hại xâm nhập vào nhà yến chủ yếu bằng đường cửa hang để bắt chim yến ăn thịt, gây giảm sút về số lượng chim, làm cho đàn yến hoảng sợ và bỏ đi nơi khác gây thiệt hại nặng cho chủ đầu tư. Giải pháp phòng ngừa thú săn chim yến hiện nay thường là đặt bẫy hoặc vây bắt. Các giải pháp này gây tận diệt cho các loài săn mồi, làm mất cân bằng tự nhiên và vi phạm luật bảo vệ tài nguyên thiên nhiên của nước ta.

Trong những năm gần đây, phương pháp phân loại bằng máy học (machine learning) đã đạt những thành tựu mới với sự ra đời của các kỹ thuật học sâu (deep learning). Các kỹ thuật về mạng học sâu vẫn đang được phát triển. Hàng năm, các cuộc thi về giải thuật nhận dạng và phân loại cho mạng học sâu được tổ chức nhằm tối ưu hóa và tìm ra các thuật toán cao cấp với khả năng giải quyết vấn đề nhanh, trên cấu hình máy tính vừa phải. Trong cuộc thi ImageNet Large-Scale Visual Recognition Challenge 2014 (ILSVRC14), tập đoàn Google đã giới thiệu một giải thuật và đạt giải nhất cuộc thi mang tên GoLeNet với sức mạnh vượt trội, đang được các nhà nghiên cứu trên thế giới ứng dụng [1]-[5].

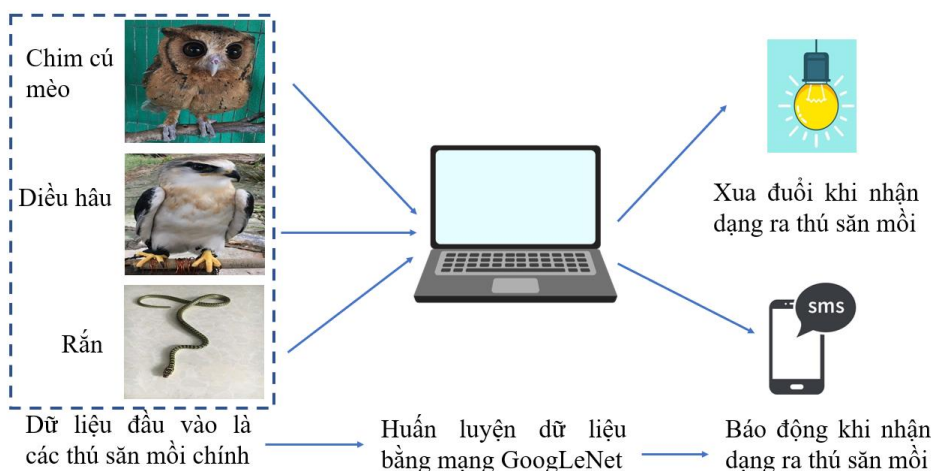
Tại Việt Nam bằng cách vận dụng kỹ thuật xử lý ảnh và học sâu (Deep Learning) nhiều công trình nghiên cứu về nhận dạng đã đạt được một số kết quả khả quan như: nhận dạng khuôn mặt dựa trên mạng nơ-ron tích chập [6], nhận dạng chữ viết tay [7], nhận dạng biển số xe [8]. Tuy nhiên, việc nghiên cứu và ứng dụng mạng học sâu GoLeNet hiện rất hạn chế. Vì vậy, nghiên cứu này đề xuất một giải pháp ứng dụng mạng học sâu GoLeNet để xây dựng mô hình thiết bị hỗ trợ xua đuổi và báo động khi có thú săn mồi xâm nhập vào nhà yến. Giải pháp đề xuất là một hệ thống bao gồm cả phần cứng và phần mềm. Phần mềm sẽ tích hợp mạng học sâu GoLeNet, được huấn luyện để phân loại và nhận dạng ra 3 loài thú săn mồi chính, đó là: chim cú mèo, điều hâu và rắn. Phần cứng được xây dựng chủ yếu dựa trên kit vi điều khiển Arduino ESP8266, đảm nhận chức năng bật đèn để xua đuổi cú mèo và điều hâu; đồng thời gửi tin nhắn SMS để báo động đến số điện thoại của người nuôi yến khi phát hiện có rắn xâm nhập.

Cấu trúc bài báo này gồm: Phần 1 giới thiệu tổng quát về vấn đề nghiên cứu, sơ lược về giải pháp được đề xuất; Phần 2 của bài báo chủ yếu giới thiệu về mạng GoLeNet, trình bày sơ lược về cơ sở dữ liệu thu thập để huấn luyện và giới thiệu về phần cứng thiết bị IoT được ứng dụng; Trong phần 3, kết quả huấn luyện mạng học sâu và thử nghiệm bước đầu sẽ được trình bày và thảo luận; Phần 4, là kết luận và đề nghị. Ngoài ra, phần phụ lục có thể hiện mã QR-Code chứa đường dẫn của clip minh họa thử nghiệm bước đầu của hệ thống này.

## 2. Phương pháp xây dựng hệ thống

### 2.1. Tổng quan về hệ thống

Kiến trúc tổng quan của hệ thống nhận dạng, xua đuổi và cảnh báo một số loài thú săn mồi thường xuyên xâm nhập vào nhà yến được mô tả như hình 1. Hệ thống đề xuất bao gồm một máy tính dùng để huấn luyện bằng phương pháp học sâu, với tập dữ liệu là ảnh 03 loài động vật săn mồi thường gặp, gồm: chim cú mèo, điều hâu và rắn. Sau khi huấn luyện và kiểm tra đạt kết quả tốt nhất, mạng học sâu sẽ được dùng để kiểm nghiệm trên các video và hình ảnh khác, được sưu tập từ nhiều nguồn trên Internet để đánh giá khả năng đáp ứng của mạng. Thiết bị IoT được đặt ngưỡng kích hoạt chức năng báo động khi kết quả nhận dạng đối tượng cho độ chính xác đạt từ 70% trở lên.



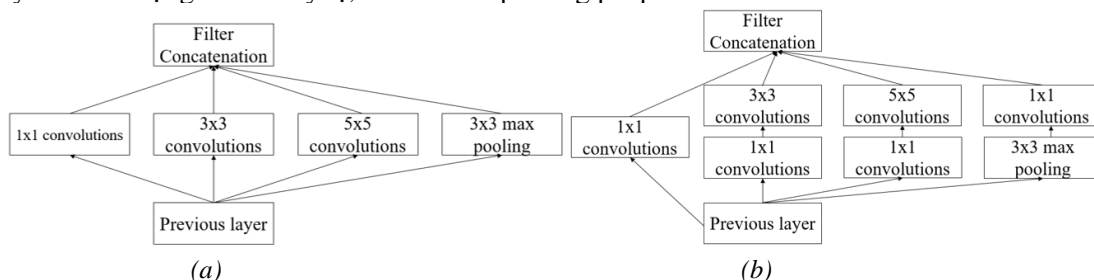
**Hình 1.** Hệ thống nhận dạng, xua đuổi và cảnh báo một số loài thú săn mồi xâm nhập nhà yến

## 2.2. Huấn luyện mạng học sâu GoogLeNet

### 2.2.1. Đặc điểm mạng học sâu GoogLeNet

GoogLeNet là một mạng nơ-ron nhân tạo với 22 lớp học sâu và 5 lớp chập (pooling layer) được tập đoàn Google kết hợp với rất nhiều viện, trường đại học nghiên cứu và đề xuất vào năm 2014 [9]. Sức mạnh chính của mạng học sâu này là các dữ liệu đầu vào không cần thiết kể một bộ tiền xử lý ảnh như: lọc ảnh và trích xuất dữ liệu đặc trưng [10]. Kích thước của các ảnh huấn luyện và nhận dạng đều được đặt ở chuẩn 224x224x3 điểm ảnh (pixels). Theo [11], đặc điểm chính của mạng học sâu GoogLeNet gồm các lớp tích chập (convolution layer), các lớp tổng hợp tối đa (max pooling layer) và trọng tâm là các lớp khởi đầu (inception layer). Trong kiến trúc của lớp khởi đầu, kích thước của lớp tích chập là cố định.

Theo đó, lớp tích chập 1x1, 3x3, 5x5 và lớp tổng hợp tối đa (max pooling) 3x3 được thực hiện theo cách song song ở đầu vào và đầu ra của chúng, được xếp chồng lên nhau để tạo ra đầu ra cuối cùng (hình 2). Với phương pháp này, các bộ lọc tích chập có các kích thước khác nhau sẽ xử lý các đối tượng ở nhiều tỷ lệ, tốt hơn các phương pháp trước đó.



**Hình 2.** Mô hình lớp khởi đầu của mạng (a) phiên bản đầu tiên, (b) khởi giảm kích thước

### 2.2.2. Kiến trúc mạng GoogLeNet

Kiến trúc mạng học sâu GoogLeNet gồm 2 cấu trúc cơ bản: thứ nhất là một mạng học sâu nhiều lớp (multilayer network) và thứ hai là một cấu trúc mạng – trong - mạng (network – in - network). Một mạng học sâu nhiều lớp được trình bày tại [11]. Theo đó, giả sử ta có một mạng học sâu  $D(l, \rho, l, \{G_i\})$  với xác suất cao giống nhau (các đồ thị ngẫu nhiên giữa các lớp) cho 2 nút bất kỳ  $u, v$  trong một lớp  $h(1)$ :

$$\Pr \left[ h_u^{(1)} = h_v^{(1)} = 1 \right] \begin{cases} \geq \frac{\rho_2}{2}; & \text{nếu } u, v \text{ có liên quan} \\ \leq \frac{\rho_2}{4}; & \text{các trường hợp còn lại} \end{cases} \quad (1)$$

Khi  $u, v$  có liên quan, các phép tích chập  $1 \times 1$  được tính toán nhằm tạo các bộ lọc cho bộ học sâu. Trong trường hợp này, mô hình tích chập  $1 \times 1$  được trình bày tại [12] với đặc điểm cấu trúc lớp – theo – lớp. Mô hình này bao gồm một đơn vị hiệu chỉnh tuyến tính:

$$f_{i,j,k} = \max(\omega_k^T x_{i,j}, 0) \quad (2)$$

Với  $(i, j)$  là chỉ số điểm ảnh trong bản đồ đối tượng (pixel index),  $x_{ij}$  là viết tắt của đặc điểm đầu vào trập trung tại vị trí  $(i, j)$  và  $k$  là giá trị để lập vị trí các kênh của bản đồ đối tượng.

Tiếp theo đó, một lớp mạng học sâu được gọi là *mlpconv* được tính toán dựa theo phương trình:

$$\begin{aligned} f_{i,j,k_1}^1 &= \max(\omega_{k_1}^1 T x_{i,j} + b_{k_1}, 0) \\ &\vdots \\ f_{i,j,k_n}^n &= \max(\omega_{k_n}^n T f_{i,j}^{n-1} + b_{k_n}, 0) \end{aligned} \quad (3)$$

với  $n$  là số lượng lớp của perceptron nhiều lớp và bộ đơn vị hiệu chỉnh tuyến tính như (2), được sử dụng như một hàm kích hoạt của các perceptron nhiều lớp.

Như vậy, mô hình GoogLeNet hoạt động theo quy tắc là khi tạo một lớp tiếp theo trong mô hình học sâu, ta chú ý đến kết quả của lớp trước. Đặc điểm này được thiết kế để tính toán hiệu quả hơn cho các máy tính có cấu hình không quá mạnh. Mạng học sâu GoogLeNet gồm những yếu tố cơ bản sau đây:

- Lớp gộp trung bình có kích thước bộ lọc  $5 \times 5$  và độ dài 3, kết quả là một lớp  $4 \times 4 \times 512$  hoặc  $4 \times 4 \times 528$ .
- Một tích chập  $1 \times 1$  với 128 bộ lọc để giảm kích thước và kết hợp với một bộ hiệu chỉnh tuyến tính (rectified linear unit).
- Một lớp được kết nối đầy đủ 1024 lớp và một bộ nội suy tuyến tính.
- Một lớp loại bỏ, với tỉ lệ loại bỏ 70% ở ngõ ra.
- Lớp tuyến tính sử dụng hàm tổn thất softmax để phân loại (dự đoán 1000 lớp giống như bộ phân loại chính, nhưng bị loại bỏ tại thời điểm suy luận).

### 2.3. Xây dựng tập dữ liệu

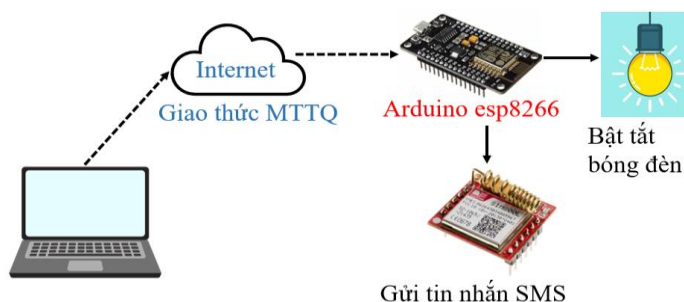
Trong nghiên cứu này, tập dữ liệu được thu thập từ 3 đối tượng săn mồi chính trong nhà yến gồm rắn, điều hâu và cú mèo. Camera quan sát được thiết kế ngay cửa vào của nhà yến và tại các vị trí cố định bên trong, hướng ra ngoài cửa. Tại cửa ra vào, chim yến bay với tốc độ cao và không đậu lại nên camera thông thường không thể ghi hình kịp. Tuy nhiên, đối với động vật săn mồi, chúng có đặc điểm là đáp (bò) xuống bệ cửa, đứng yên quan sát và chọn lựa con mồi thích hợp. Do vậy, việc dùng camera để ghi hình và chụp ảnh là khả thi. Từ những cơ sở trên, tập dữ liệu dùng để huấn luyện mạng học sâu GoogLeNet được xây dựng trên bảng 1. Ngoài ra, nghiên cứu này còn chọn lọc và bổ sung một số ảnh thu thập được từ nhiều nguồn khác nhau trên mạng Internet.

**Bảng 1.** Số lượng ảnh và nhãn được chuẩn bị để huấn luyện

Nhãn	Số ảnh huấn luyện	Số ảnh kiểm chứng
Rắn	2000	521
Điều hâu	600	320
Cú mèo	1800	500

### 2.4. Xây dựng phần cứng

Phần cứng thiết bị được xây dựng dựa theo yêu cầu của chủ nhà yến. Cụ thể là khi phát hiện động vật săn mồi, hệ thống sẽ xua đuổi và cảnh báo cho gia chủ. Cấu trúc chung của các nhà yến là bên trong rất tối nên các loài chim săn như điều hâu và cú mèo rất sợ ánh sáng được phát lên bất ngờ từ đèn; trong khi đó, cách duy nhất để đối phó với rắn là trực tiếp bắt để chuyển đi nơi khác. Vì vậy, một bộ điều khiển tự động từ máy tính được xây dựng để bật đèn xua đuổi chim cú mèo và điều hâu, đồng thời gửi tin nhắn SMS tới điện thoại di động của chủ nhà khi phát hiện có rắn xâm nhập. Từ những yêu cầu trên, một hệ phần cứng được thiết kế theo hình 3.



**Hình 3.** Mô hình phần cứng cảnh báo thú săn môi

Ngay khi mạng học sâu GoogLeNet nhận dạng được thú săn môi với độ chính xác lớn hơn 70%, máy tính sẽ điều khiển phần cứng bật sáng đèn để xua đuổi hoặc gửi SMS cảnh báo. Trọng tâm trong hệ phần cứng này là vi điều khiển Arduino ESP8266, giao tiếp với máy tính thông qua mạng Internet bởi giao thức MQTT (Message Queuing Telemetry Transport).

Giao thức MQTT lần đầu được giới thiệu vào năm 1999 như một hệ thống cung cấp (publish) và thuê bao (subscribe) được sử dụng cho các thiết bị IoT, với băng thông thấp. Trung gian thuê bao/cung cấp là một MQTT Borker, về cơ bản, đó là một server trung gian để lưu giữ tín hiệu truyền lên từ thiết bị gửi và gửi đến thiết bị nhận. Hệ thống này có thể bảo vệ dữ liệu thông qua một token, được tạo ngẫu nhiên mật mã, góp phần bảo vệ hệ thống. Trong mô hình này, máy tính đóng vai trò như một đơn vị cung cấp và vi điều khiển Arduino ESP8266 đóng vai trò như một thuê bao.

### 3. Kết quả và thảo luận

#### 3.1. Kết quả huấn luyện GoogLeNet

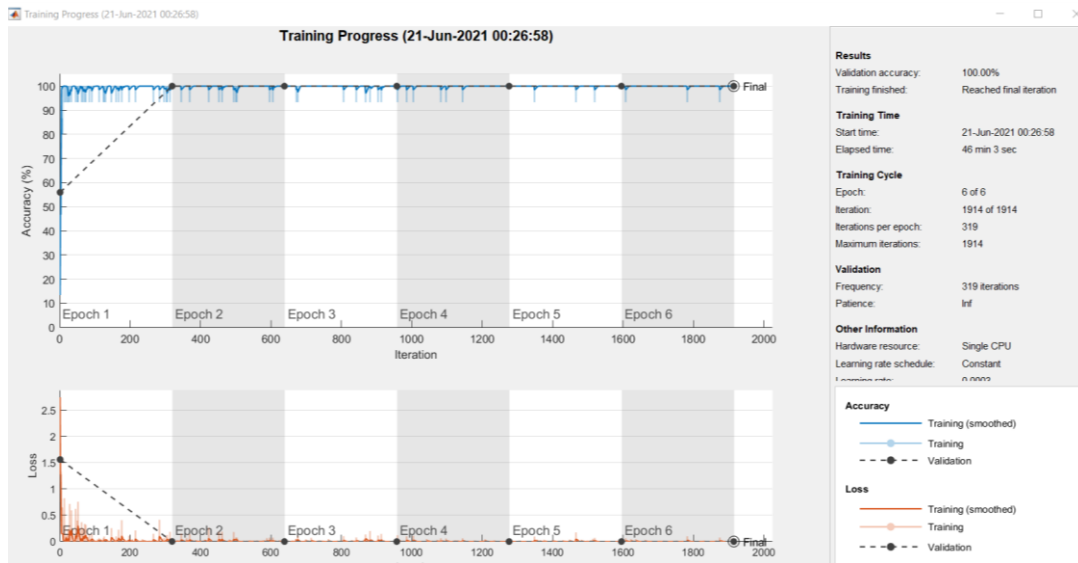
Ngôn ngữ lập trình cấp cao MATLAB được sử dụng để huấn luyện và nhận dạng cho mạng học sâu GoogLeNet. MATLAB được chạy trên máy tính HP Prodesk core i5-9600 RAM 16GB, windows 10 64 bit, để huấn luyện mạng. Một tập hợp các tùy chọn để huấn luyện cần được tạo ra. Vì vậy, hệ thống phân tích, học sâu với phương pháp giảm gradient ngẫu nhiên động lượng, gồm 144 lớp. Theo đó, hệ thống có 9 lớp khởi đầu (inception). Ngõ ra của hệ thống là một tập dữ liệu được đưa tiếp vào mạng huấn luyện mạng - trong - mạng (network - in - network), để huấn luyện tiếp tục. Số chu kỳ học tối đa bằng 6.

Hình 4 miêu tả quá trình huấn luyện của mạng GoogLeNet. Tập dữ liệu được sử dụng để huấn luyện mạng, từ bảng 1, tương ứng với cột 'Số ảnh huấn luyện'. Tuy nhiên, chỉ 80% số ảnh này được đưa vào huấn luyện mạng và 20% còn lại dùng để kiểm tra ngay trong quá trình huấn luyện (validation). Thời gian huấn luyện của mạng khoảng 48 phút với 6 chu kỳ học cho cả 2 quá trình tăng cường chính xác và suy giảm. Toàn bộ quá trình bao gồm 1,914 lần lặp lại với 319 lần lặp lại cho mỗi chu kỳ học. Trên biểu đồ hình 4, độ chính xác khi mạng ước lượng trên dữ liệu kiểm tra (validation data) tăng từ gần 70% lên 100% ở chu kỳ đầu và giữ nguyên như vậy tới hết 6 lần lặp lại. Độ chính xác khi mạng đánh giá trên dữ liệu huấn luyện (training data) và làm mượt dao động đạt hơn 90%. Tại biểu đồ hình 4, tỷ lệ sai số đối với dữ liệu kiểm tra giảm từ 1,5% về gần 0% trong chu kỳ học đầu và ổn định tới hết quá trình huấn luyện.

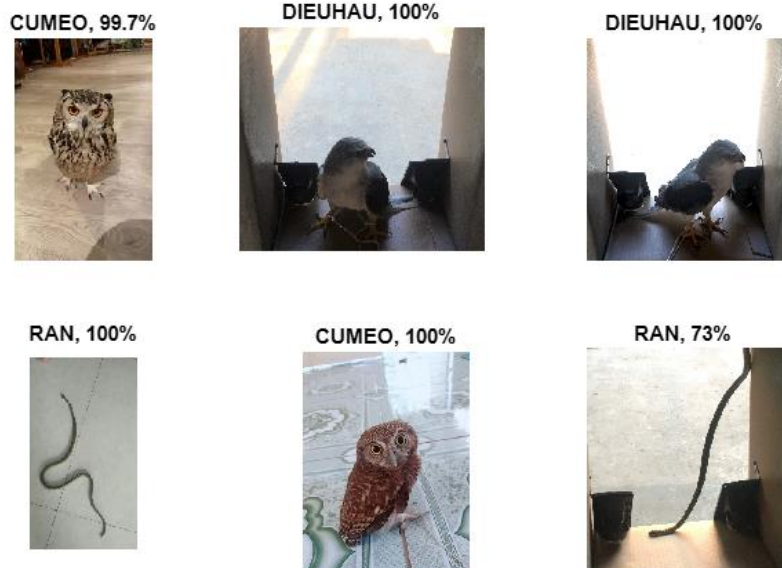
Kết thúc quá trình huấn luyện, mạng GoogLeNet sẽ được kiểm chứng lại để đánh giá độ chính xác dựa theo bộ dữ liệu tương ứng ở cột 'Số ảnh kiểm chứng' trong bảng 1.

#### 3.2. Nhận dạng đối tượng gây hại bằng mạng học sâu GoogLeNet

Hình 5 minh họa kết quả nhận dạng của mạng học sâu GoogLeNet, trên tập dữ liệu kiểm chứng, với hình ảnh thu thập được từ nhiều nguồn khác nhau. Các kết quả trên hình 5 cho thấy, mạng học sâu có thể nhận dạng tốt 2 loài chim cú và điều hâu với độ chính xác gần như 100%. Đối với rắn, do cách di chuyển đặc biệt, trong một số trường hợp độ chính xác chỉ đạt 73%, tuy nhiên trong trường hợp này, tín hiệu báo động cũng đã được kích hoạt.



Hình 4. Quá trình huấn luyện của mạng học sâu GoogLeNet



Hình 5. Kết quả phân loại bằng GoogLeNet

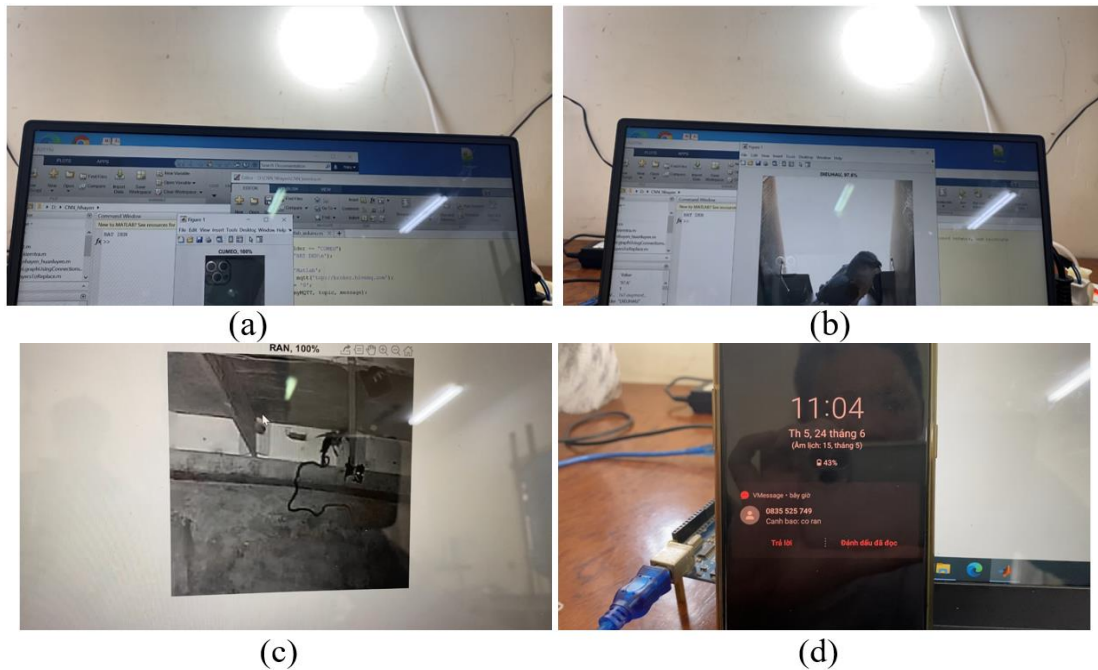
Kết quả thử nghiệm mạng học sâu GoogLeNet trên toàn bộ tập dữ liệu kiểm chứng được thể hiện ở bảng 2. Theo đó, tỉ lệ nhận dạng đúng cao nhất ở cú mèo, đạt 99,6% và thấp nhất ở rắn, đạt được 90,8%. Ngoài ra, nghiên cứu này cũng đã thu thập từ mạng Internet 3 clips được quay trong nhà yến, có xuất hiện thú săn mồi, gồm: 2 clips có chim cú và 1 clip có rắn xâm nhập. Để kiểm tra khả năng đáp ứng của giải thuật, hình ảnh từ các clips này được trích xuất và kiểm tra thêm khả năng nhận dạng của mạng. Kết quả minh họa trên hình 6. Theo đó, việc nhận dạng cú mèo và rắn cho kết quả rất tốt, với độ chính xác gần như tuyệt đối (~100%).

Bảng 2. Thử nghiệm khả năng phân loại với nhiều nguồn ảnh khác nhau

Nhãn	Số ảnh thử nghiệm	Số ảnh nhận diện đúng	Tỷ lệ (%)
Rắn	521	473	90,8
Điều hâu	320	316	98,7
Cú mèo	500	498	99,6



**Hình 6.** Kết quả phát hiện thú săn mồi từ clip trong nhà yến



**Hình 7.** Kết quả hoạt động của mô hình  
 (a) *Bật đèn khi nhận ra cú mèo;* (b) *Bật đèn khi nhận ra điều hâu*  
 (c) *Hệ thống phát hiện ra rắn;* (d) *Gửi tin nhắn SMS cảnh báo*

Từ các thử nghiệm trên, việc kết hợp phần mềm và phần cứng đã được triển khai. Kết quả thử nghiệm được trình bày trên hình 7. Clip thử nghiệm có link tham khảo theo mã QR-Code trong phần phụ lục. Nhìn chung hệ thống đáp ứng tốt yêu cầu đề ra, đèn được bật sáng ngay khi kết quả nhận dạng có độ chính xác lớn hơn ngưỡng cảnh báo (>70%).

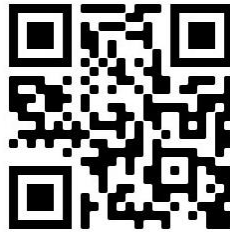
#### 4. Kết luận và đề nghị

Nghiên cứu này đã đề xuất và thử nghiệm một mô hình phát hiện, xua đuổi và cảnh báo khi có động vật săn mồi xâm nhập nhà nuôi chim yến. Mạng học sâu GoogLeNet được sử dụng để huấn luyện và phân loại động vật săn mồi. Kết quả nhận dạng cho thấy, cú mèo và điều hâu có thể nhận dạng với tỉ lệ chính xác cao, trên 90%; trong khi đó, rắn có tỉ lệ nhận dạng thấp hơn, do cách thức di chuyển dễ tương đồng với môi trường. Phần cứng được điều khiển không dây thông qua vi điều khiển Arduino ESP8266, có thể bật tắt đèn và gửi SMS cảnh báo đúng thời điểm.

Thời gian tới, nhóm nghiên cứu sẽ tiếp tục bổ sung thêm dữ liệu cho 3 loài hiện hữu, đồng thời xây dựng thêm cơ sở dữ liệu cho một số loài động vật săn mồi khác như: tắc kè và dơi mặt

quỹ, nhằm đáp ứng yêu cầu vừa bảo vệ chim yến, vừa không cần tiêu diệt các loài động vật khác. Việc lắp đặt hệ thống vào nhà yến thực tế đã thỏa thuận được với các đối tác tình nguyện và sẽ sớm có kết quả ở một báo cáo khác.

#### Phụ lục clip minh họa



Clip hoạt động của toàn bộ mô hình hệ thống kèm chú thích

#### TÀI LIỆU THAM KHẢO/ REFERENCES

- [1] J. Huang, Q. Zhang, and L. Qin, "Adapted GooLeNet for Answer Selection," *2nd IEEE Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC)*, 2018, pp. 1256-1262.
- [2] J. Huang, Y. Hu, and W. Yang, "Adapted GooLeNet for Visual Question Answering," *3rd International Conference on Mechanical, Control and Computer Engineering (ICMCCE)*, 2018, pp. 603-606.
- [3] J. Ma, J. Rao, Y. Qiao, and W. Liu, "Sprouting Potato Recognition Based on Deep Neural Network GooLeNet," *IEEE 3rd International Conference on Cloud Computing and Internet of Things (CCIOT)*, 2018, pp. 502-505.
- [4] J. Huang, Q. Zhang, and L. Qin, "Adapted GooLeNet for Answer Selection," *2nd IEEE Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC)*, 2018, pp. 1256-1262.
- [5] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Image Net classification with deep convolutional neural networks," *Advances in Neural Information Processing Systems*, 2012, pp. 1097-1105.
- [6] H. Q. Doan, H. M. Le, and N. T. Doan, "Face recognition in video using convolutional neural network," *Vietnam science and technology*, no. 1, pp. 8-12, 2020.
- [7] P. Q. Pham and Q. P. Vuong, "Recognition of handwriting digits using artificial neuron network," *Hue University Journal of Science*, no. 1, pp. 119-129, 2019.
- [8] S. P. Ho, V. D. Phan, V. C. Le, and H. C. Ta, "Applying neural networks, convolutional neural networks and combination of cpus and gpus to increase calculation performance for image classification," *Vinh University Journal of science*, vol. 47, pp. 64-76, 2018.
- [9] C. Szegedy, L. Wei, J. Yangqing, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going Deeper with Convolutions," *ArXiv:1409.4842 [Cs]*, Sept. 2014.
- [10] B. Zhou, A. Khosla, A. Lapedriza, A. Torralba, and A. Oliva, "Places: An image database for deep scene understanding," *arXiv preprint arXiv:1610.02055*, 2016.
- [11] S. Arora, A. Bhaskara, R. Ge, and T. Ma "Provable Bounds for Learning Some Deep Representations," *ArXiv:1310.6343 [Cs, Stat]*, arXiv.org, Oct. 2013.
- [12] L. Min, C. Qiang, and Y. Shuicheng, "Network In Network," *ArXiv:1312.4400 [Cs]*, arXiv.org, Mar. 2014.